

Best Available Copy

0151593

日 本 国 特 許 庁

JAPAN PATENT OFFICE

1017 U.S. PRO
10/083563
02/27/02

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 7月30日

出 願 番 号

Application Number:

特願2001-229500

出 願 人

Applicant(s):

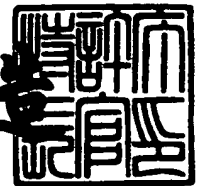
株式会社東芝

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年12月14日

特 許 庁 長 官
Commissioner,
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3108936



特 2 0 0 1 - 2 2 9 5 0 0

【書類名】 特許願

【整理番号】 A000103849

【提出日】 平成13年 7月30日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 9/44

【発明の名称】 知識分析システムならびに同システムにおける分析条件
設定方法、分析条件保存方法および再分析処理方法

【請求項の数】 14

【発明者】

【住所又は居所】 東京都青梅市末広町2丁目9番地 株式会社東芝青梅工
場内

【氏名】 永村 栄治

【発明者】

【住所又は居所】 東京都青梅市末広町2丁目9番地 株式会社東芝青梅工
場内

【氏名】 近藤 修明

【発明者】

【住所又は居所】 東京都青梅市末広町2丁目9番地 株式会社東芝青梅工
場内

【氏名】 高知尾 勝彦

【発明者】

【住所又は居所】 東京都青梅市末広町2丁目9番地 株式会社東芝青梅工
場内

【氏名】 渥美 一彦

【発明者】

【住所又は居所】 東京都青梅市末広町2丁目9番地 株式会社東芝青梅工
場内

【氏名】 佐々木 淳哉

【発明者】

【住所又は居所】 東京都青梅市末広町2丁目9番地 株式会社東芝青梅工場内

【氏名】 島川 和典

【特許出願人】

【識別番号】 000003078

【氏名又は名称】 株式会社 東芝

【代理人】

【識別番号】 100058479

【弁理士】

【氏名又は名称】 鈴江 武彦

【電話番号】 03-3502-3181

【選任した代理人】

【識別番号】 100084618

【弁理士】

【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 知識分析システムならびに同システムにおける分析条件設定方法、分析条件保存方法および再分析処理方法

【特許請求の範囲】

【請求項 1】 複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムであって、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御手段と、

前記アクセス制御手段によってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識をカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析手段とを具備し、

前記知識分析手段は、

前記クラスタデータベースの作成時に、クラスタリングで優先する重要語を設定する手段を有し、前記重要語をもとにクラスタの軸を生成するようにクラスタリングを実行することを特徴とする知識分析システム。

【請求項 2】 前記知識分析手段は、更にクラスタリングで無視する不要語を設定する手段を有し、前記重要語と前記不要語をもとにクラスタリングを実行することを特徴とする請求項 1 記載の知識システム。

【請求項 3】 前記知識分析手段は、更にクラスタリングで同じ単語として扱う同義語を設定する手段を有し、前記重要語と前記不要語および前記同義語の組合せによってクラスタリングを実行することを特徴とする請求項 2 記載の知識分析システム。

【請求項 4】 複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムであって、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御手段と、

前記アクセス制御手段によってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識を重要語、不要語、同義語の分析条件をもとにカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析手段とを具備し、

前記知識分析手段は、

前記クラスタデータベースの作成時に用いた前記分析条件を保存する手段を有することを特徴とする知識分析システム。

【請求項 5】 前記知識分析手段は、前記保存した分析条件を読み出し、変更を施した後の分析条件を用いて、新たなクラスタデータベースを作成する手段を有することを特徴とする請求項 4 記載の知識分析システム。

【請求項 6】 複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムであって、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御手段と、

前記アクセス制御手段によってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識を重要語、不要語、同義語の分析条件をもとにカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析手段とを具備し、

前記知識分析手段は、

前記クラスタデータベースの作成時に用いた前記分析条件を取得し、前記重要語、前記不要語、前記同義語の少なくとも 1 つ以上を再設定した後の分析条件を用いて、クラスタデータベースを再作成して置換する手段を有することを特徴とする知識分析システム。

【請求項 7】 複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムにおける分析条件設定方法であって、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御ステップと、

前記アクセス制御ステップによってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識をカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析ステップとを具備し、

前記知識分析ステップは、

前記クラスタデータベースの作成時に、クラスタリングで優先する重要語を設定する分析条件設定ステップを有することを特徴とする知識分析システムにおける分析条件設定方法。

【請求項 8】 前記知識分析ステップは、更にクラスタリングで無視する不要語を設定するステップを有し、前記重要語と前記不要語をもとにクラスタリングを実行することを特徴とする請求項 7 記載の知識システムにおける分析条件設定方法。

【請求項 9】 前記知識分析ステップは、更にクラスタリングで同じ単語として扱う同義語を設定するステップを有し、前記重要語と前記不要語および前記同義語の組合せによってクラスタリングを実行することを特徴とする請求項 8 記載の知識分析システムにおける分析条件設定方法。

【請求項 10】 複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムにおける分析条件保存方法であって、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御ステップと、

前記アクセス制御ステップによってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識を重要語、不要語、同義語の分析条件をもとにカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析ステップとを具備し、

前記知識分析ステップは、

前記クラスタデータベースの作成時に用いた前記分析条件を保存する分析条件保存ステップを有することを特徴とする知識分析システムにおける分析条件保存

方法。

【請求項 11】 複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムにおける再分析処理方法であって、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御ステップと、

前記アクセス制御ステップによってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識を重要語、不要語、同義語の分析条件をもとにカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析ステップとを具備し、

前記知識分析ステップは、

前記クラスタデータベースの作成時に用いた前記分析条件を取得し、前記重要語、前記不要語、前記同義語の少なくとも1つ以上を再設定した後の分析条件を用いて、クラスタデータベースを再作成して置換する再分析処理ステップを有することを特徴とする知識分析システムにおける再分析処理方法。

【請求項 12】 複数のクライアント端末とネットワークを介して接続可能に構成されたサーバコンピュータに、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムにおける分析条件設定機能を実現させるためのプログラムであって、

前記サーバコンピュータを、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御手段と、

前記アクセス制御手段によってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識をカテゴリ毎に定義されるクラスタ別に分類するためのクラスタデータベースを作成する知識分析手段として機能させ、

前記知識分析手段に、

前記クラスタデータベースの作成時に、クラスタリングで優先する重要語、ク

ラスタリングで無視する不要語およびクラスタリングで同じ単語として扱う同義語を設定する分析条件設定手順を実行させることを特徴とするプログラム。

【請求項13】 複数のクライアント端末とネットワークを介して接続可能に構成されたサーバコンピュータに、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムにおける分析条件保存機能を実現させるためのプログラムであって、

前記サーバコンピュータを、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御手段と、

前記アクセス制御手段によってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識を重要語、不要語、同義語の分析条件をもとにカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析手段として機能させ、

前記知識分析手段に、

前記クラスタデータベースの作成時に用いた前記分析条件を保存する分析条件保存手順を実行させることを特徴とするプログラム。

【請求項14】 複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムにおける再分析処理機能を実現させるためのプログラムであって、

前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御手段と、

前記アクセス制御手段によってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識を重要語、不要語、同義語の分析条件をもとにカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析手段として機能させ、

前記知識分析手段に、

前記クラスタデータベースの作成時に用いた前記分析条件を取得し、前記重要語、前記不要語、前記同義語の少なくとも1つ以上を再設定した後の分析条件を

用いて、クラスタデータベースを再作成して置換する再分析処理手順を実行させることを特徴とするプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、ナレッジマネジメントシステムで用いられる知識分析システムならびに同システムにおける分析条件設定方法、分析条件保存方法および再分析処理方法に係り、特に、クラスタの軸となる重要語およびこの重要語と組み合わせてクラスタリングを実行させるための不要語および同義語を設定可能とするとともに、この重要語、不要語、同義語を含む分析条件を保存して再クラスタリング時に利用することを可能とした知識分析システムならびに同システムにおける分析条件設定方法、分析条件保存方法および再分析処理方法に関する。

【0002】

【従来の技術】

近年、企業を中心に複数のユーザ間で情報共有を行うためのグループウェアの導入が進められている。代表的なグループウェアとしては、電子メールシステムやワークフローシステムなどが知られているが、最近では、知識や情報の共有支援を図るためのナレッジマネジメントシステムも開発され始めている。

【0003】

このナレッジマネジメントシステムは、Web情報や電子ファイル情報などに加え、個人のノウハウなどを知識データベースとして蓄積・管理するためのものであり、自然言語検索などの検索機能と組み合わせることにより、知識、情報の効率的な活用が可能となる。

【0004】

ところで、このようなナレッジマネジメントシステムにおいては、個人のノウハウなどの知識をどのように収集・蓄積するかが重要となる。個人のノウハウなどの知識はいわゆる暗黙知であって、Web情報や電子ファイル情報などのように形式化されたものではないため、それを自動的に収集、蓄積することは困難であるからである。

【0005】

そこで、最近では、知識蓄積支援機能を持つナレッジマネジメントシステムの開発が要求されている。個人のノウハウなどの知識を自動的に収集・蓄積する仕組みを実現することにより、暗黙知としての知識をもWeb情報や電子ファイル情報などのような形式化された形式知と同様に活用することが可能となる。

【0006】

また、このようにして蓄積された知識や情報を容易に検索するナレッジマネジメントシステムの開発も並行して行われている。典型例としては、自然言語の質問文を入力して、有用な知識や情報を検索する知識検索支援機能を持つ自然言語検索システムが挙げられる。

【0007】

【発明が解決しようとする課題】

ところで、この種のナレッジマネジメントシステムでは、知識や情報を簡単に整理または閲覧したり、あるいは、初めて利用するユーザに対しても、どのような知識情報が検索できるのかをわかりやすく提示する等、知識データの有効活用を図るための知識の体系化が強く求められる。これに伴い、蓄積した知識をカテゴリ化し階層化することにより分類すること等が行われている。

【0008】

大量に蓄積された知識の分類は、非常に手間の掛かる作業であるため、たとえば形態素解析等の技術を利用して自動化することが望ましい。しかしながら、自動的に分類を行わせると、その分類時のカテゴリ化および階層化が、その時点で蓄積された知識群の内容的な偏りに大きく左右されてしまい、必ずしもユーザを満足させる結果が得られないといった問題があった。

【0009】

この発明は、このような事情を考慮してなされたものであり、クラスタの軸を生成する重要語およびこの重要語と組み合わせてクラスタリングを実行させるための不要語および同義語を設定可能とするとともに、この重要語、不要語、同義語を含む分析条件を保存して再クラスタリング時に利用することを可能とした知識分析システムならびに同システムにおける分析条件設定方法、分析条件保存方

法および再分析処理方法を提供することを目的とする。

【0010】

【課題を解決するための手段】

前述した目的を達成するために、この発明は、複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムであって、前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御手段と、前記アクセス制御手段によってアクセス許可された各クライアント端末からの知識分析に用いるクラスターデータベースであって、前記知識データベースに蓄積された知識をカテゴリ毎に定義されるクラスター別にクラスターデータベースを作成する知識分析手段とを具備し、前記知識分析手段は、前記クラスターデータベースの作成時に、クラスターリングで優先する重要語を設定する手段を有し、前記重要語をもとにクラスターの軸を生成するようにクラスターリングを実行することを特徴とする知識分析システムを提供する。

【0011】

また、前記知識分析手段は、更にクラスターリングで無視する不要語を設定する手段を有し、前記重要語と前記不要語をもとにクラスターリングを実行することを特徴とする。

【0012】

また、前記知識分析手段は、更にクラスターリングで同じ単語として扱う同義語を設定する手段を有し、前記重要語と前記不要語および前記同義語の組合せによってクラスターリングを実行することを特徴とする。

【0013】

この知識分析システムにおいては、クラスターの軸を生成する重要語やこの重要語と組み合わせてクラスターリングを行わせるための不要語および同義語を設定する仕組みをもつことにより、その時点で蓄積された知識群の内容的な偏りに左右されることなく、ユーザが意図するカテゴリ化および階層化によって分類を実行することを可能とする。

【 0 0 1 4 】

また、この発明は、複数のクライアント端末とネットワークを介して接続可能に構成され、知識データベースに蓄積された知識に対する各クライアント端末からの分析を支援する知識分析システムであって、前記クライアント端末からの知識分析を許可するために、アクセス要求元のクライアント端末に対してユーザ認証を行うアクセス制御手段と、前記アクセス制御手段によってアクセス許可された各クライアント端末からの知識分析に用いるクラスタデータベースであって、前記知識データベースに蓄積された知識を重要語、不要語、同義語の分析条件をもとにカテゴリ毎に定義されるクラスタ別にクラスタデータベースを作成する知識分析手段とを具備し、前記知識分析手段は、前記クラスタデータベースの作成時に用いた前記分析条件を保存する手段を有することを特徴とする知識分析システムを提供する。

【 0 0 1 5 】

また、前記知識分析手段は、前記保存した分析条件を読み出し、変更を施した後の分析条件を用いて、新たなクラスタデータベースを作成する手段を有することを特徴とする。

【 0 0 1 6 】

この知識分析システムにおいては、クラスタリング実行時の分析条件を保存しておき、この保存した分析条件を呼び出し、必要に応じて、重要語、不要語および同義語を再設定して再クラスタリングを実行できるようにすることにより、より精度の高い分析を求めることを可能とし、かつ、再分析の操作性および効率を飛躍的に向上させる。

【 0 0 1 7 】

【発明の実施の形態】

以下、図面を参照してこの発明の一実施形態を説明する。

【 0 0 1 8 】

図 1 には、この発明の実施形態に係る知識分析システムの構成が示されている。この知識分析システムは、知識分析機能を持つナレッジマネジメントシステムとして利用されるものであり、複数のクライアント端末 1 1 が共通に利用可能な

知識データベース 1 2 2 5 を用いて、知識の分析を行う。具体的な構成の説明を行う前に、まず、図 2 乃至図 9 を用いて、この実施形態に係る知識分析システムの概要を説明することにする。

【 0 0 1 9 】

この実施形態の知識分析システムは、ある目的で集められた雑多な文書群、つまり知識群から読み取れる傾向等の把握を支援するシステムである。この知識分析システムは次の 3 つの機能から構成される。

【 0 0 2 0 】

(1) 知識クラスタの自動作成

類似する知識を集め、知識クラスタ（知識のグループ）に自動的に仕分け、編集する機能。

(2) 知識クラスタ一覧の表示

仕分けられた知識クラスタ間での知識数などを比較する機能。

(3) 知識クラスタ内知識の時系列グラフ表示

ある 1 つの知識クラスタでの知識の発生傾向を見る機能。

これらにより

- ・ アンケート等の分析
- ・ ヘルプデスクへの質問の傾向分析
- ・ アイディアの分類、

営業日報からの売れ筋商品や要因分析

掲示板からの最新話題の抽出

思わぬ発見／新たな気づき（新たな知識の創造）

などといった応用に役立てることができる。

【 0 0 2 1 】

以下、それぞれの機能説明をする。

【 0 0 2 2 】

< 知識クラスタの自動作成 >

図 2 に、知識クラスタの自動作成概念を、図 3 に、クラスタリングの流れを示す。ナレッジマネジメントシステムのデータベースに集められた大量の知識を、

同じような語が使われているかどうかをもとに、似た内容どうしのグループに仕分けて、知識クラスタ（知識グループ）を作成する。また、クラスタの中にクラスタがある、という階層として仕分けることもでき、さらに、この作成した知識クラスタ群を「分類」として名前（分類名）をつけて保存することもできる。

【0023】

（1）検索時の指定

検索時の指定には、以下の2つが存在する。

【0024】

- ・ 検索条件：キーワード検索。
- ・ クラスタリング対象とする知識数（検索スコアの上位いくつまでを使うか）。

（2）指定事項A

指定項目Aには、以下の9つが存在する。

【0025】

- ・ 重要語／不要語／同義語の指定。
- ・ 重要語：分析に重要と想定される単語。クラスタリングで優先される。
- ・ 不要語：分析に不必要な単語。クラスタリングで無視される。
- ・ 同義語：分析で同じ意味を持つ単語群。クラスタリングで同じ単語として扱われる。

【0026】

- ・ 作成する知識クラスタの個数（第1階層）。
- ・ 階層を1つとするか、階層数を特に指定しないか。
- ・ 知識が1つのクラスタにしか入らないか（1対1）、複数に入るか（1対n）。
- ・ ラベルを名詞のみとするか他の品詞も混ぜるか。
- ・ クラスタ全体の名前（分類名）。

（3）指定事項B（再クラスタリング時の指定）。

指定項目Bには、以下の4つが存在する。

【0027】

- ・ 固定クラスターの指定（再クラスタリング時でも壊れないクラスター）。
- ・ 統合するクラスター（2つ以上のクラスター→1つのクラスター）の指定。
- ・ 重要語／不要語の指定。
- ・ 作成する知識クラスターの個数（第1階層：再クラスタリング後にできるクラスター数）。

【0028】

図2の知識データベース1225には、多くの知識が順序付けもなくまた整理もなされずに格納されている。この知識データベース1225からクラスタリング対象としたい知識の条件を指定して検索すると（図3ステップA1）、図2のように検索された知識の集合が決定され（ステップA2）、上記の指定事項Aによりクラスタリングの際の条件を指定すると（ステップA3）、クラスタリング（自動グルーピング）が実行されてクラスター名生成、階層構造決定、知識割り当てがなされ（ステップA4）、図2の知識クラスター1～知識クラスターnが最初の知識クラスター群として作成・保存される（ステップA5、ステップA6）。

【0029】

こうして最初にクラスター群が作成されるが、エンドユーザは異なる条件で再度クラスタリングしたい場合（ステップA7のYES）、必要に応じて指定条件の調整（上記の指定事項B）を行い（ステップA8）、再クラスタリング（これをフィードバックと称する）を実行する（ステップA9）。これにより、再クラスタリングの結果として調整された知識クラスター群が作成・保存される（ステップA10、ステップA6）

一方、再クラスタリングが当初より不要な場合、または、これ以上の再クラスタリングが不要な場合は（ステップA7のNO）、このクラスタリング処理を終了する。

【0030】

<知識クラスター一覧の表示>

（1）階層非表示&グラフ非表示

表示は知識の件数順に行われる。図4は、件数順の全クラスター一覧の表示画面例を示している。階層を無視した全クラスターの一覧表示と、特定の階層（第1階

層、第2階層、…、第n階層）での一覧表示の2種類がある。以下、図4の画面の説明を行う。

【0031】

- ・ クラスタ名：クラスタ名が表示される。
- ・ 件数：クラスタに属する知識の件数が表示される。
- ・ キーワード：そのクラスタのキーワードが表示される。

(2) 階層非表示&グラフ表示

図5は、階層に関係なく全クラスタが一覧で表示された画面である。階層を無視した全クラスタの一覧表示と、特定の階層での一覧表示との2種類がある。以下、図5の画面の説明を行う。

【0032】

- ・ クラスタ名：クラスタ名が表示される。
- ・ 件数：クラスタに属する知識の件数が表示される。
- ・ キーワード：そのクラスタのキーワードが表示される。
- ・ グラフ：知識件数が棒グラフにて表示される。

【0033】

(3) 階層表示&グラフ非表示

全てのクラスタに対して、階層関係を表示する方法には次の2つがある。

【0034】

(a) 階層全体の表示

(b) 階層を固定しての表示と階層無視での表示

図6は、階層全体の画面表示例を示している。これは、分析結果が階層構造のみで表示される画面である。以下、この画面の説明を行う。

【0035】

- ・ クラスタ：分析されたクラスタと、そのクラスタのキーワードが階層構造で表示される。

【0036】

- ・ 件数：クラスタ下の知識件数が表示される。

【0037】

(4) 階層表示&グラフ表示

全てのクラスタに対して、階層関係を表示する方法には次の2つがある。

【0038】

(a) 階層全体の表示

(b) 階層を固定しての表示と階層無視での表示

図7は、階層全体の画面表示例を示している。これは、分析結果が階層構造とグラフで表示される画面である。以下、この画面の説明を行う。

【0039】

- ・クラスタ：分析されたクラスタと、そのクラスタのキーワードが階層構造で表示される。

- ・件数：クラスタ下の知識件数が表示される。

【0040】

- ・グラフ：知識件数が棒グラフにて表示される。また、階層ごとに、違う色で表示される。

<知識クラスタ内知識の時系列グラフ表示>

知識クラスタに登録された知識の数を時系列（日ごと、月ごと）でグラフ表示する。

【0041】

(1) 時系列分析表示

クラスタ下の登録知識を、登録された時間情報に従って月または日単位に表示する機能である。画面起動時は、グラフは表示されておらず、表示単位および表示範囲を指定することで指定範囲のグラフが表示される。図8は、月単位で表示した例である。以下、この画面の説明を行う。

【0042】

- ・表示単位：グラフを表示する単位を月単位または日単位のいずれかに指定する。デフォルトは「月単位」。

- ・表示範囲：表示する範囲を指定する。

- ・グラフ表示：月単位の場合、表示範囲の指定年に適合する知識の件数を月単位に集計の上、グラフ表示する。このとき、表示範囲の「月」の項目に数値を

入力してもグラフ表示には反映されない。一方、日単位の場合、表示範囲の指定年月に適合する知識の件数を日単位に集計の上、グラフ表示する。

【0043】

(2) 分類一覧

クラスタリングを行なった結果（知識クラスタの固まり）は、分類として保存される。保存された分類の一覧表示を行なう画面が分類一覧である。図9の画面が知識分析機能の初期画面となる。

【0044】

図9の画面から分類を選択し、その分類のなかのクラスター一覧を見ることが出来る。新規作成もこの画面からできる。知識分析機能の起動時、この図9の画面が表示され、分析結果を選択する画面となる。以下、この画面の説明を行う。

【0045】

- ・ 分析結果：分析した結果の名称が表示される。この名称は分析結果の作成時にユーザが指定する。

- ・ 知識DB：分析に使用した知識DBの名称が表示される。

- ・ 件数：分析したDBの全件数が表示される。

- ・ 最上位クラスタ数：第1階層のクラスタ数が表示される。

- ・ 更新日：分析結果を最後に更新した日付が表示される。

- ・ 作成日：分析結果を作成した日付が表示される。

<システム構成>

次に、図1を参照して、この実施形態に係る知識分析システムのシステム構成について説明する。

【0046】

この実施形態の知識分析システムは、複数のクライアント端末11とLAN等のコンピュータネットワーク13を介して相互に接続可能なサーバコンピュータ12にて実現されている。サーバコンピュータ12とクライアント端末11とは、それぞれ、図示しないが、CPU、メインメモリ、記憶装置としての磁気ディスク装置およびキーボードやマウスなどの入力部とディスプレイなどの表示部とを持つ入出力装置が設けられている。

【0047】

クライアント端末11では、Webブラウザ111が動作している。サーバコンピュータ12上に構築された知識分析のためのリソースを示すURL (Uniform Resource Locator) をWebブラウザ111から指定することにより、知識分析処理を各クライアント端末11から利用することができる。

【0048】

サーバコンピュータ12の知識分析機能は、主に、Webサーバ121の制御モジュール1211、ナレッジサーバ122の登録モジュール1221、検索モジュール1222およびクラスタリングモジュール1223などのソフトウェアと、これらソフトウェアによって知識分析のために利用される管理情報および実データとによって実現されている。管理情報には、各クライアント端末11に対してユーザ認証を行うためのログイン管理情報1212が存在する。また、実データとしては、文書データベース1224、知識データベース1225、中間処理用データベース1226および分析結果データベース1227が存在する。

【0049】

制御モジュール1211は、知識分析に関する全体の動作を制御するためのものであり、この知識分析システムの中核プログラムであるナレッジサーバ122とWebサーバ121との間の仲介機能を初め、Webサーバ121を通じて各クライアント端末11がナレッジサーバ122にログインする際のユーザ認証機能を持つ。このユーザ認証のために、制御モジュール1211は、ログイン管理情報1212を管理している。このログイン管理情報1212には、この知識分析システムに参加しているユーザそれぞれのユーザIDとパスワード等が格納されている。このユーザ認証により、各クライアント端末11からの知識分析等の為になされるナレッジサーバ122に対するアクセスの許可・禁止の制御が行われる。

【0050】

ナレッジサーバ122は、複数のクライアント端末11が分析可能な知識となる知識データベース1225や分析結果データベース1227の管理、運用を行うためのものであり、各クライアント端末11から指定された条件による知識分

析結果を知識のクラスタとして分類、蓄積する。

【0051】

ここで、ナレッジサーバ122のクラスタリングモジュール1223によって実行される、この知識分析システムにおける分析条件設定処理、分析条件保存処理および再分析処理について説明する。分析条件設定処理とは、クラスタの軸となる重要語およびこの重要語と組み合わせてクラスタリングを実行させるための不要語および同義語を設定する機能であり、また、分析条件保存処理は、この重要語、不要語および同義語を含む分析条件を保存する機能である。そして、再分析処理は、保存された分析条件を用いて効率的に知識分析を再実行するための機能である。

【0052】

まず、図10乃至図15の説明を行う。図10は、知識データベース選択画面である。図10には以下の項目が表示される。

【0053】

・知識DB（コメント／件数）：知識データベース1225の名称、コメントおよび件数。

なお、知識データベース1225は、文書データベース1224と一対に設けられるものであり、文書データベース1224に蓄積される文書、つまり知識に含まれる語句の索引を格納するものである。クライアント端末11からいずれかの文書データベース1224に対する文書の格納を依頼されると、ナレッジサーバ122の登録モジュール1221は、この文書を指定された文書データベース1224に格納するが、その際、この文書の内容に対してたとえば形態素解析を施し、その文書に含まれる語句の索引を作成する。そして、その作成した索引を知識データベース1225に格納する。

【0054】

また、図11は、図10で選択した知識データベース1225を分析する条件を指定する画面である。図11には、以下の項目が表示される。

【0055】

・分析対象DB：図10の知識データベース選択画面で選択したすべての知識

データの名称、コメントおよび 全件数を表示する。

- ・分析結果名称：分析した結果を保存する際の名称を入力する。本名称がすでに保存してある分析結果と重複する場合は、＜実行＞選択時に警告メッセージが表示され、そのメッセージ中の＜保存＞を選択すると、従来の分析結果に上書きして保存される。

- ・分析対象期間：分析対象となる期間を入力する。入力のない場合は全知識が分析対象になる。

- ・絞込キーワード：分析のキーワードとなる語句を入力する。入力のない場合は、全知識が分析対象になる。

- ・絞込件数：分析結果に登録される最大件数を入力する。分析結果が分析件数よりも大きくなった場合は、分析条件との適合度が高い方から指定件数分の知識が抽出される。入力のない場合は、分析条件に適合するすべての知識が抽出される。

- ・階層数：クラスタリングを行う階層数を指定する。デフォルトは「多階層」。

- ・知識の重複：1つの知識が複数のクラスタに重複して登録されることを許可するかどうかを指定する。「あり」を選択した場合、各知識は関連のあるクラスタすべてに登録され、その結果、同じ知識が別の複数のクラスタ内に存在することになる。このため、分析結果の合計は、分析元DBの全件数の和を上回る場合がある。一方、「なし」を選択した場合、各知識は最も関連の強いクラスタ1つにのみ登録される。デフォルトは「あり」。

- ・最上位クラスタの最大個数：最上位の階層に作成されるクラスタの最大個数を指定する。入力のない場合は、指定なしとしてクラスタリングを行う。

【0056】

- ・重要語：分析を行う上での重要語を入力する。複数入力する際は、スペース区切りで行う。

- ・不要語：分析を行う上での不要語を入力する。複数入力する際は、スペース区切りで行う。

- ・同義語：分析を行う上での同義語を入力する。同義語同士を“=”で結んで

記述し、セミコロン（；）で区切る。1項目に3つ以上の同義語を定義することも可能である。また、定義した同義語は、分析結果では、一番左に記述された単語に集約されることになる。

【0057】

また、図12は、図11で指定した知識データベース1225の分析条件を保存する画面である。図12には、以下の項目が表示される。

【0058】

- ・分析条件名：すべての分析条件

また、図13は、図12で保存した知識データベース1225の分析条件を読み出す画面である。図13には、以下の項目が表示される。

【0059】

- ・分析条件名
- ・作成日時

また、図14は、図6のクラスター一覧画面（クラスタ毎の件数のグラフ表示画面に同じ）で編集ボタンがクリックされたときに表示される画面であり、再分析の対象となるクラスタ名を選択する画面である。図11には、以下の項目が表示される。

【0060】

- ・固定：そのクラスタが固定クラスタかどうかを表示する。固定クラスタの場合、再分析を実行してもそのクラスタ下は変更されない。

- ・クラスタ名：クラスタ名を表示する。
- ・キーワード：そのクラスタのキーワードを表示する。
- ・件数：クラスタに所属する文書の件数を表示する。

さらに、図15は、図14のクラスタ編集画面で再分析ボタンがクリックされたときに表示される画面であり、再分析条件を指定する画面である。つまり、すでに作成されている分析結果であるクラスタを再分析して上書きするための分析条件を入力する画面である。図15には、以下の項目が表示される。なお、分析条件の各フィールドには、前回の分析で用いられたパラメータが入力された状態で表示される。このうち、最上位クラスタの最大個数と重要語、不要語および同

義語のみが再分析で変更できる。

【0061】

- ・分析対象DB：分析に使用した知識データベース1225の名称が表示される。
- ・分析結果名称：分析した結果が保存されている名称を表示する。
- ・分析対象期間：分析対象となる期間が表示される。
- ・絞込キーワード：分析のキーワードとなる語句が表示される。
- ・絞込件数：分析結果に登録される最大件数が表示される。
- ・階層数：クラスタリングを行う階層数を表示する。
- ・知識の重複：1つの知識が複数のクラスタに重複して登録されることを許可するかどうかを表示する。
- ・最上位クラスタの最大個数：最上位の階層に作成されるクラスタの最大個数を指定する。入力のない場合は、指定なしとしてクラスタリングを行う。変更可能。

【0062】

- ・重要語：分析を行う上での重要語を入力する。複数入力する際は、スペース区切りで行う。

【0063】

- ・不要語：分析を行う上での不要語を入力する。複数入力する際は、スペース区切りで行う。

【0064】

- ・同義語：分析を行う上での同義語を入力する。同義語同士を“=”で結んで記述し、セミコロン（;）で区切る。1項目に3つ以上の同義語を定義することも可能である。また、定義した同義語は、分析結果では、一番左に記述された単語に集約されることになる。

【0065】

図16は、図10で選択した知識データベース1225に対して図11の条件設定画面で指定した重要語、不要語および同義語に基づく知識分析を実行し、その際に指定した重要語、不要語および同義語を含む分析条件を保存する手順を示

すフローチャートである。

【 0 0 6 6 】

ユーザがWebブラウザ111によりサーバコンピュータ12の制御モジュール1211へログイン要求すると(ステップB1)、制御モジュール1211は、ユーザから入力されたユーザIDおよびパスワードが登録されているか否かを調べるためにログイン管理情報1212にアクセスし(ステップB2)、このログインを許可するかどうかを判定するためのユーザ認証を行う(ステップB3)。ユーザIDおよびパスワードがログイン管理情報1212に登録されておらず、ログインが失敗した場合(ステップB3のNO)、制御モジュール1211は、ログイン失敗をWebサーバ121を通じてWebブラウザ111に返してこの処理を終了する(ステップB4)。

【 0 0 6 7 】

一方、ユーザIDおよびパスワードがログイン管理情報1212に登録されており、ログインが成功した場合には(ステップB3のYES)、ナレッジサーバ122のクラスタリングモジュール1223が、図10に示す知識データベース選択画面を表示する(ステップB5)。この知識データベース選択画面が表示されると、ユーザは、その画面上で分析対象の知識データベース1225を選択する(ステップB6)。この選択が行われると、ナレッジサーバ122のクラスタリングモジュール1223は、今度は図11に示す分析条件指定画面を表示する(ステップB7)。

【 0 0 6 8 】

また、この分析条件指定画面が表示されると、ユーザは、この画面上で分析条件を指定する(ステップB8)。この設定が行われると、ナレッジサーバ122の検索モジュール1222が、知識データベース選択画面で選択された知識データベース1225に対し、分析条件指定画面で指定された重要語、不要語および同義語を用いた検索を行い、中間処理用データベース1226を作成する(ステップB9)。次に、ナレッジサーバ122のクラスタリングモジュール1223が、分析条件指定画面で指定された分析条件に基づき、クラスタリング処理を実行する(ステップB10)。その後、ユーザからの指示に従い、その分析条件を

保存する場合には（ステップB11のYES）、図12に示す分析条件保存画面で分析条件を保存した上でこの処理を終了する（ステップB12）。一方、保存しない場合には（ステップB11のNO）、そのままこの処理を終了する。

【0069】

図17は、図10で選択した知識データベース1225に対し、図11の分析条件指定画面で分析条件読み込みボタンをクリックすることにより表示される分析条件読み込み画面（図13）において分析条件名を選択し、それが読み込まれた後に表示される図11の条件指定画面の分析条件を用いて新たにクラスタリングする場合の処理の手順を示すフローチャートである。なお、この分析条件の読み込みによるクラスタリング処理では、既存のクラスタを上書きすることは行わず、異なったクラスタが新たに作成される。そのため、読み込んだ後に分析結果名を変更する。

【0070】

ユーザがWebブラウザ111によりサーバコンピュータ12の制御モジュール1211へログイン要求すると（ステップC1）、制御モジュール1211は、ユーザから入力されたユーザIDおよびパスワードが登録されているか否かを調べるためにログイン管理情報1212にアクセスし（ステップC2）、このログインを許可するかどうかを判定するためのユーザ認証を行う（ステップC3）。ユーザIDおよびパスワードがログイン管理情報1212に登録されておらず、ログインが失敗した場合（ステップC3のNO）、制御モジュール1211は、ログイン失敗をWebサーバ121を通じてWebブラウザ111に返してこの処理を終了する（ステップC4）。

【0071】

一方、ユーザIDおよびパスワードがログイン管理情報1212に登録されており、ログインが成功した場合には（ステップC3のYES）、ナレッジサーバ122のクラスタリングモジュール1223が、図10に示す知識データベース選択画面を表示する（ステップC6）。この知識データベース選択画面が表示されると、ユーザは、その画面上で分析対象の知識データベース1225を選択する（ステップC6）。この選択が行われると、ナレッジサーバ122のクラスタ

リングモジュール1223は、今度は図11に示す分析条件指定画面を表示する（ステップC7）。

【0072】

また、この分析条件指定画面が表示されると、ユーザは、この画面上で分析条件読み込みボタンをクリックする（ステップC8）。これに応じて、ナレッジサーバ122のクラスタリングモジュール1223が、図13に示す分析条件読み込み画面を表示し（ステップC9）、ユーザが、この画面上で所望の分析条件を選択する（ステップC10）。

【0073】

この選択が行われると、ナレッジサーバ122のクラスタリングモジュール1223が、先のクラスタリング時に保存した分析条件を埋め込んだ分析条件指定画面を表示するので、ユーザは、分析結果名を変更して新たに異なるクラスタを作成するために分析実行ボタンをクリックする。すると、ナレッジサーバ122の検索モジュール1222が、知識データベース選択画面で選択された知識データベース1225から分析条件指定画面で指定された重要語、不要語および同義語を用いた検索を行い、中間処理用データベース1226を作成する（ステップC11）。そして、ナレッジサーバ122のクラスタリングモジュール1223が、分析条件指定画面で指定された分析条件に基づき、クラスタリング処理を実行する（ステップC12）。

【0074】

図18は、図6で選択したクラスタに対し、図14の画面で再分析ボタンをクリックすることにより表示される再分析条件指定画面（図15）において最上位クラスタの最大個数と重要語、不要語および同義語のいずれかを再設定して新たに再分析（再クラスタリングして既存のクラスタと置き換える）する場合の処理の手順を示すフローチャートである。

【0075】

ユーザがWebブラウザ111によりサーバコンピュータ12の制御モジュール1211へログイン要求すると（ステップD1）、制御モジュール1211は、ユーザから入力されたユーザIDおよびパスワードが登録されているか否かを

調べるためにログイン管理情報1212にアクセスし（ステップD2）、このログインを許可するかどうかを判定するためのユーザ認証を行う（ステップD3）。ユーザIDおよびパスワードがログイン管理情報1212に登録されておらず、ログインが失敗した場合（ステップD3のNO）、制御モジュール1211は、ログイン失敗をWebサーバ121を通じてWebブラウザ111に返してこの処理を終了する（ステップD4）。

【0076】

一方、ユーザIDおよびパスワードがログイン管理情報1212に登録されており、ログインが成功した場合には（ステップD3のYES）、ナレッジサーバ122のクラスタリングモジュール1223が、図6に示すクラスター一覧画面を表示する（ステップD5）。このクラスター一覧画面が表示されると、ユーザは、その画面上で再分析対象のクラスタを選択して選択ボタンをクリックする（ステップD6）。この選択が行われると、ナレッジサーバ122のクラスタリングモジュール1223は、今度は図14に示すクラスタ編集画面を表示する（ステップD7）。

【0077】

また、このクラスタ編集画面が表示されると、ユーザは、この画面上で再分析ボタンをクリックする（ステップD8）。これに応じて、ナレッジサーバ122のクラスタリングモジュール1223が、図15に示す再分析条件指定画面を表示し（ステップD9）、ユーザが、この画面上でたとえば重要語として新たに「金融再生」と「不良債権」とを追加して再分析ボタンをクリックする（ステップD10）。

【0078】

この分析条件の再指定が行われると、ナレッジサーバ122の検索モジュール1222が、その重要語、不要語および同義語を用いた検索を行い、中間処理用データベース1226を作成する（ステップD11）。そして、ナレッジサーバ122のクラスタリングモジュール1223が、図15で指定されたクラスタリングに関する条件に基づき、クラスタリング処理を実行する（ステップD12）。

【0079】

このように、この知識分析システムは、クラスタの軸を生成する重要語やこの重要語と組み合わせてクラスタリングを行わせるための不要語および同義語を設定する仕組みをもつことにより、その時点で蓄積された知識群の内容的な偏りに左右されることなく、ユーザが意図するカテゴリ化および階層化によって分類を実行することを可能とする。

【0080】

また、クラスタリング実行時の分析条件を保存しておき、この保存した分析条件を呼び出し、必要に応じて、重要語、不要語および同義語を再設定して再クラスタリングを実行できるようにすることにより、より精度の高い分析を求めることを可能とし、かつ、再分析の操作性および効率を飛躍的に向上させる。

【0081】

そして、このような構成をもつことにより、この知識分析システムは、図19に示すような処理の流れを簡単に作り出すことが可能となる。

【0082】

すなわち、導入時には、重要語や不要語、同義語を設定した上で初期分類階層構造の自動生成(1)を実行した後、この重要語、不要語および同義語を再設定して再クラスタリング、つまりカテゴリと知識の編集(2)を納得がいくまで実行することにより、ほぼ理想に近い分類を実現する。また、運用時には、保存しておいた分析条件を用いた再クラスタリングをたとえば定期的に行うことにより、新着文書(知識)の自動分類(3)を行い、必要に応じて、この重要語、不要語および同義語を再設定して再クラスタリングを実行することにより、カテゴリの細分化(4)を行う。

【0083】

このように、この知識分析システムは、導入時および運用時双方の操作を大幅に効率化することを可能とする。

【0084】

なお、本願発明は、前記実施形態に限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で種々に変形することが可能である。更に、前記実施形

態には種々の段階の発明が含まれており、開示される複数の構成要件における適宜な組み合わせにより種々の発明が抽出され得る。たとえば、実施形態に示される全構成要件から幾つかの構成要件が削除されても、発明が解決しようとする課題の欄で述べた課題が解決でき、発明の効果の欄で述べられている効果が得られる場合には、この構成要件が削除された構成が発明として抽出され得る。

【0085】

【発明の効果】

以上、詳述したように、この発明によれば、クラスタの軸を生成する重要語やこの重要語と組み合わせてクラスタリングを行わせるための不要語および同義語を設定する仕組みをもつことにより、その時点で蓄積された知識群の内容的な偏りに左右されることなく、ユーザが意図するカテゴリ化および階層化によって分類を実行することを可能とする。

【0086】

また、クラスタリング実行時の分析条件を保存しておき、この保存した分析条件を呼び出し、必要に応じて、重要語、不要語および同義語を再設定して再クラスタリングを実行できるようにすることにより、より精度の高い分析を求めることを可能とし、かつ、再分析の操作性および効率を飛躍的に向上させる。

【図面の簡単な説明】

【図1】

この発明の実施形態に係る知識分析システムのシステム構成を示すブロック図。

【図2】

同実施形態の知識分析システムにおける知識クラスタの自動作成概念を説明するための図。

【図3】

同実施形態の知識分析システムにおけるクラスタリングの流れを示すフローチャート。

【図4】

同実施形態の知識分析システムにおけるクラスター一覧の表示画面（階層非表示

& グラフ非表示) の一例を示す図。

【図 5】

同実施形態の知識分析システムにおけるクラスター一覧の表示画面（階層非表示 & グラフ表示）の一例を示す図。

【図 6】

同実施形態の知識分析システムにおけるクラスタ階層関係全体の表示画面（階層表示 & グラフ非表示）の一例を示す図。

【図 7】

同実施形態の知識分析システムにおけるクラスタ階層関係全体の表示画面（階層表示 & グラフ表示）の一例を示す図。

【図 8】

同実施形態の知識分析システムにおける各クラスタの時系列表示画面の一例を示す図。

【図 9】

同実施形態の知識分析システムにおける分類一覧の表示画面の一例を示す図。

【図 1 0】

同実施形態の知識分析システムにおける知識データベース選択画面の一例を示す図。

【図 1 1】

同実施形態の知識分析システムにおける分析条件指定画面の一例を示す図。

【図 1 2】

同実施形態の知識分析システムにおける分析条件保存画面の一例を示す図。

【図 1 3】

同実施形態の知識分析システムにおける分析条件読み込み画面の一例を示す図。

【図 1 4】

同実施形態の知識分析システムにおけるクラスタ編集画面の一例を示す図。

【図 1 5】

同実施形態の知識分析システムにおける再分析条件指定画面の一例を示す図。

【図 16】

同実施形態の知識分析システムにおいて、知識分析に設定した重要語、不要語および同義語を含む分析条件を保存する手順を示すフローチャート。

【図 17】

同実施形態の知識分析システムにおいて、条件設定画面の分析条件を用いて新たにクラスタリングする場合の処理の手順を示すフローチャート。

【図 18】

同実施形態の知識分析システムにおいて、再分析条件指定画面において最上位クラスターの最大個数と重要語、不要語および同義語のいずれかを再指定して新たに再分析する場合の処理の手順を示すフローチャート。

【図 19】

同実施形態の知識分析システムにより実現される知識分析の処理の流れを示すための図。

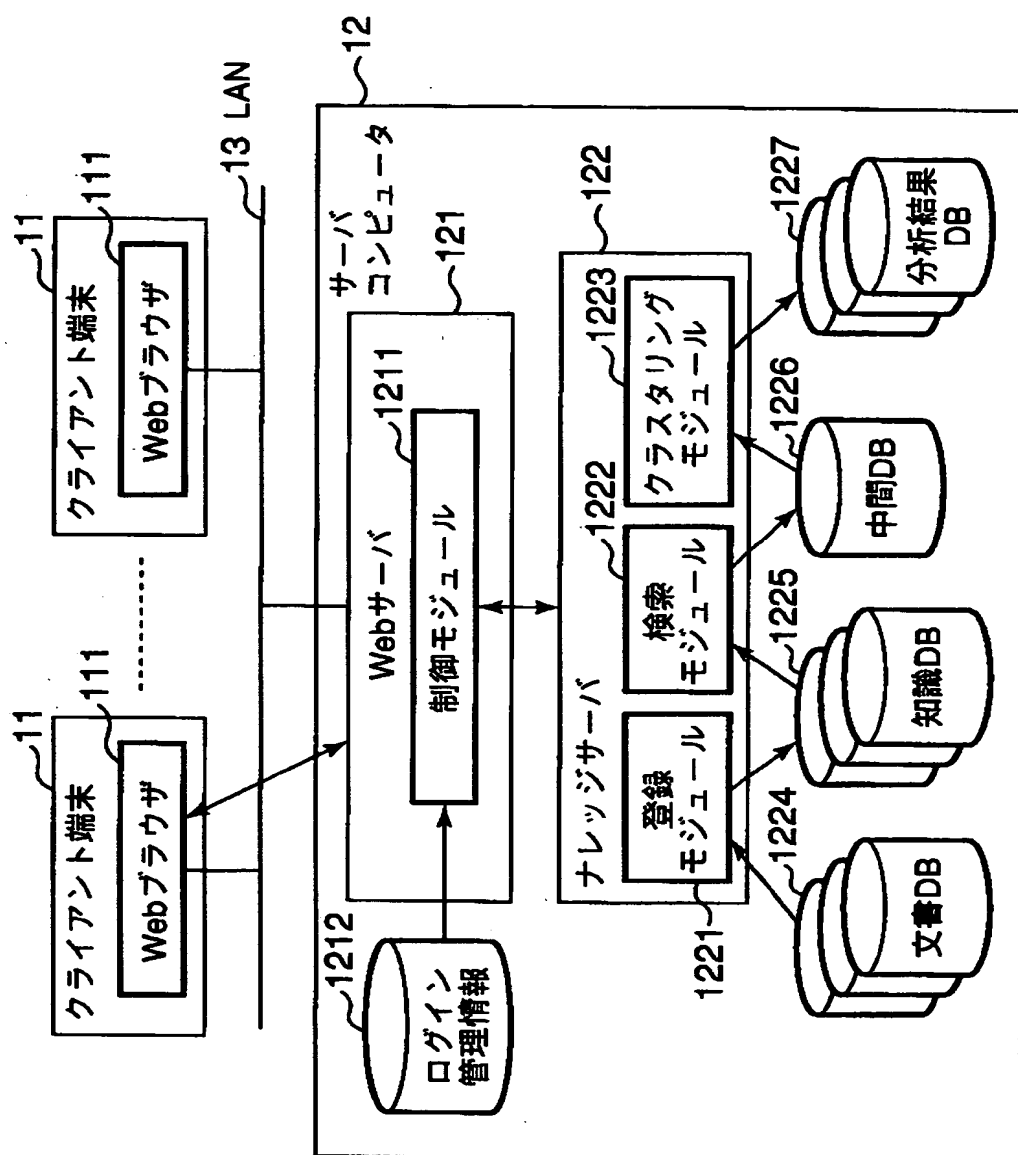
【符号の説明】

- 1 1 …クライアント端末
- 1 2 …サーバコンピュータ
- 1 3 …LAN
- 1 1 1 …Webブラウザ
- 1 2 1 …Webサーバ
- 1 2 2 …ナレッジサーバ
- 1 2 1 1 …制御モジュール
- 1 2 1 2 …ログイン管理情報
- 1 2 2 1 …登録モジュール
- 1 2 2 2 …検索モジュール
- 1 2 2 3 …クラスタリングモジュール
- 1 2 2 4 …文書データベース
- 1 2 2 5 …知識データベース
- 1 2 2 6 …中間データベース
- 1 2 2 7 …分析結果データベース

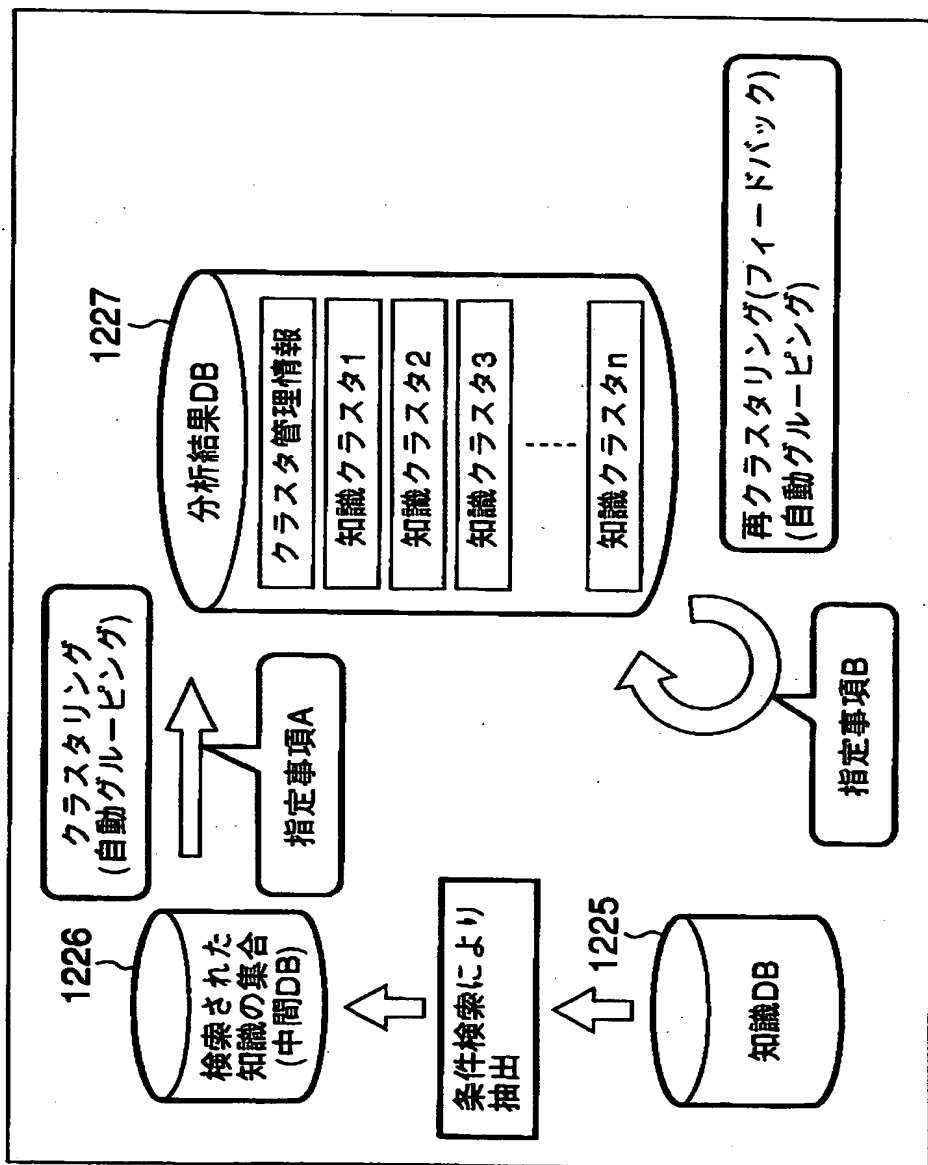
【書類名】

図面

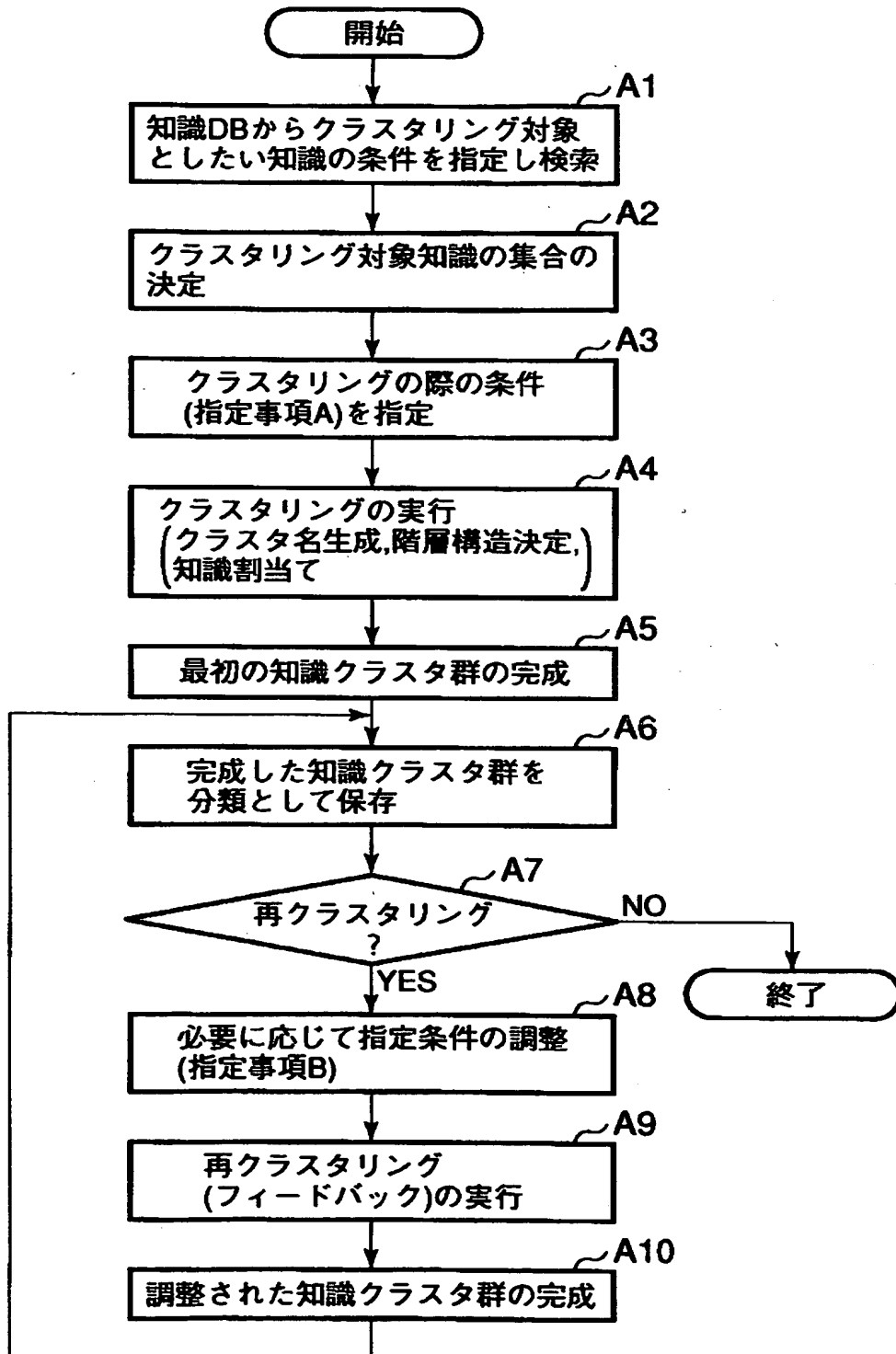
【図 1】



【図2】



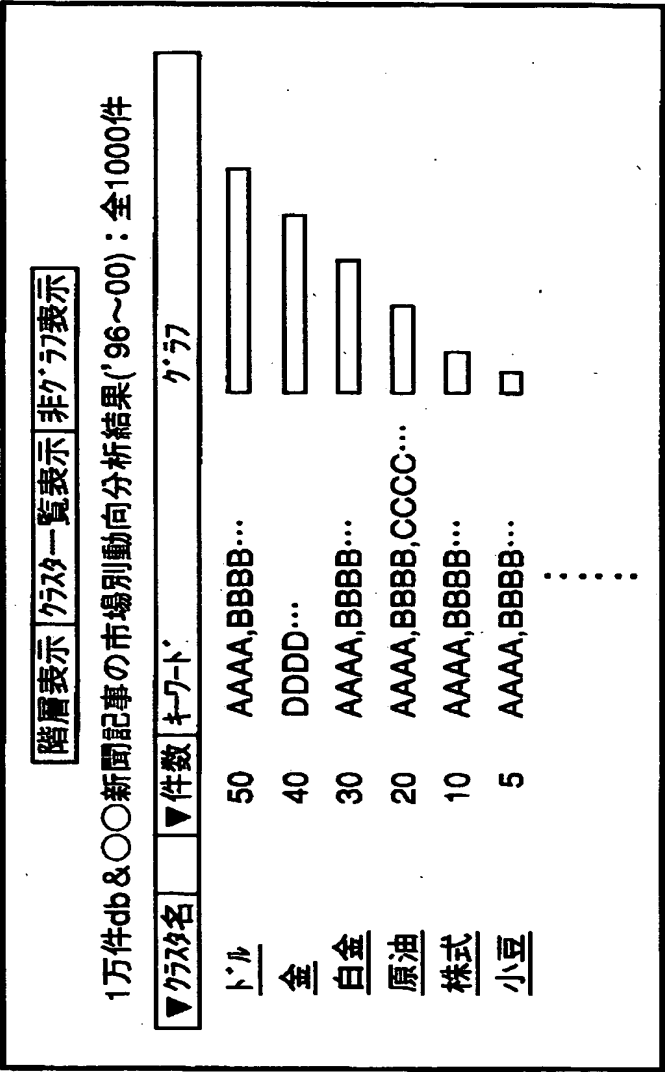
【図 3】



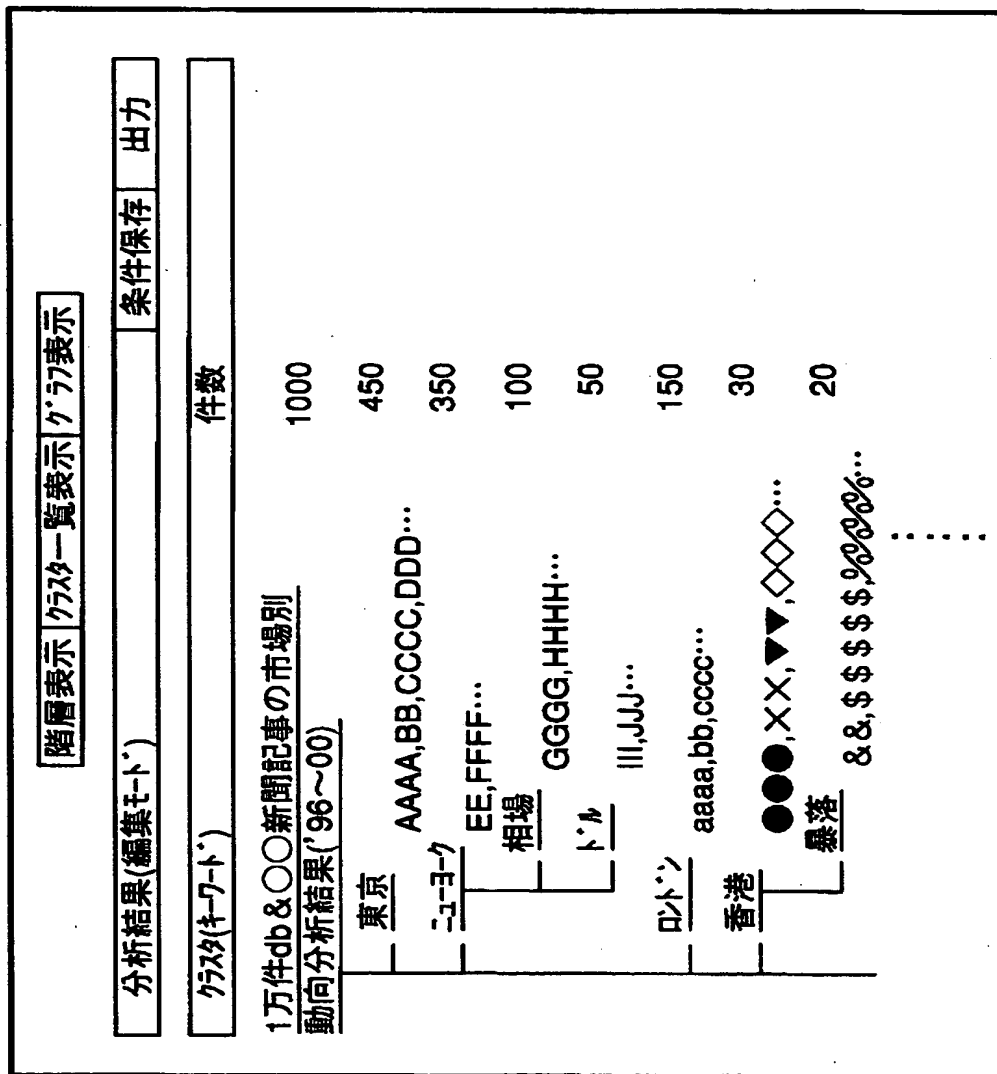
【図 4】

階層表示 クラスター表示 クラフ表示		
1万件db & ○○新聞記事の市場別動向分析結果('96~00) : 全1000件		
▼クラス名	▼件数	ネットワーク
<u>ドル</u>	50	AAAA,BBBB...
<u>金</u>	40	DDDD...
<u>白金</u>	30	AAAA,BBBB...
<u>原油</u>	20	AAAA,BBBB,CCCC...
<u>株式</u>	10	AAAA,BBBB...
<u>小豆</u>	5	AAAA,BBBB...
		...

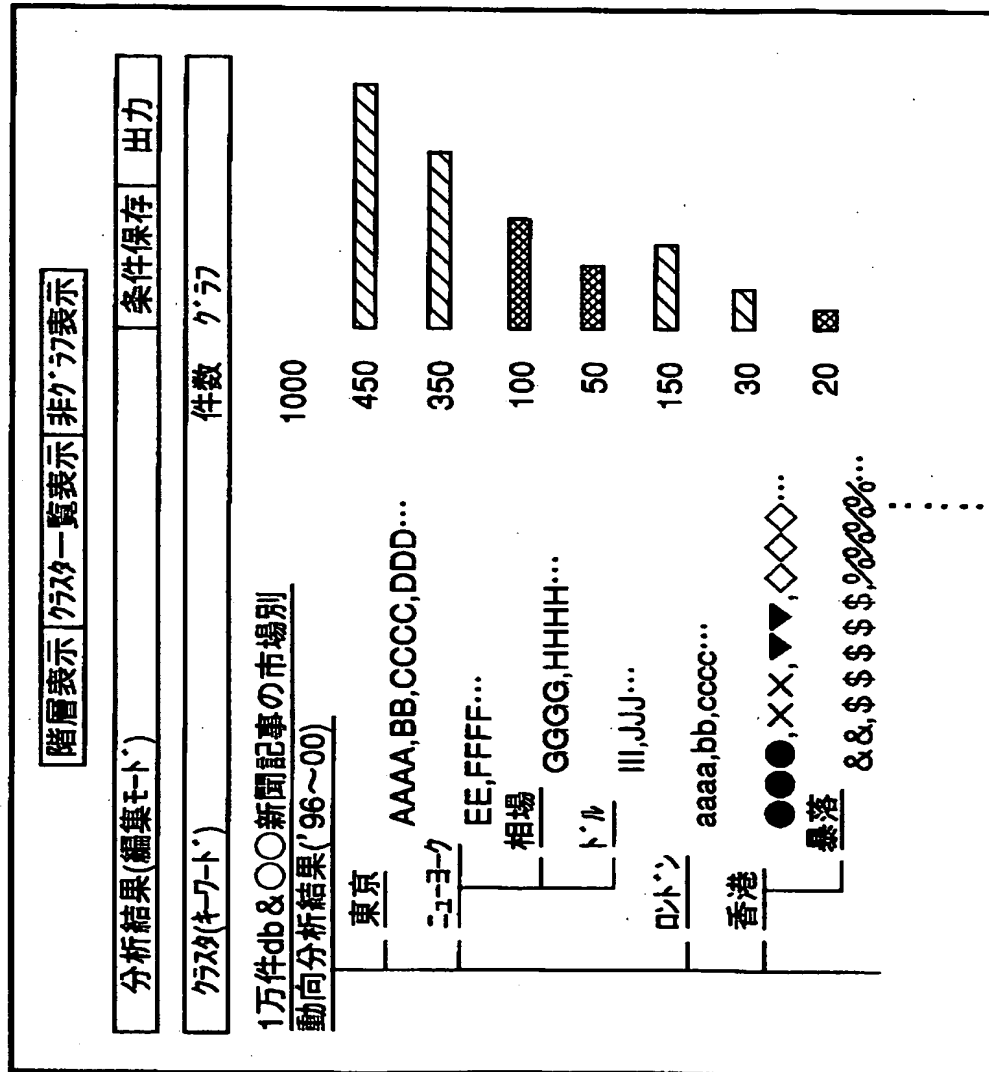
【図5】



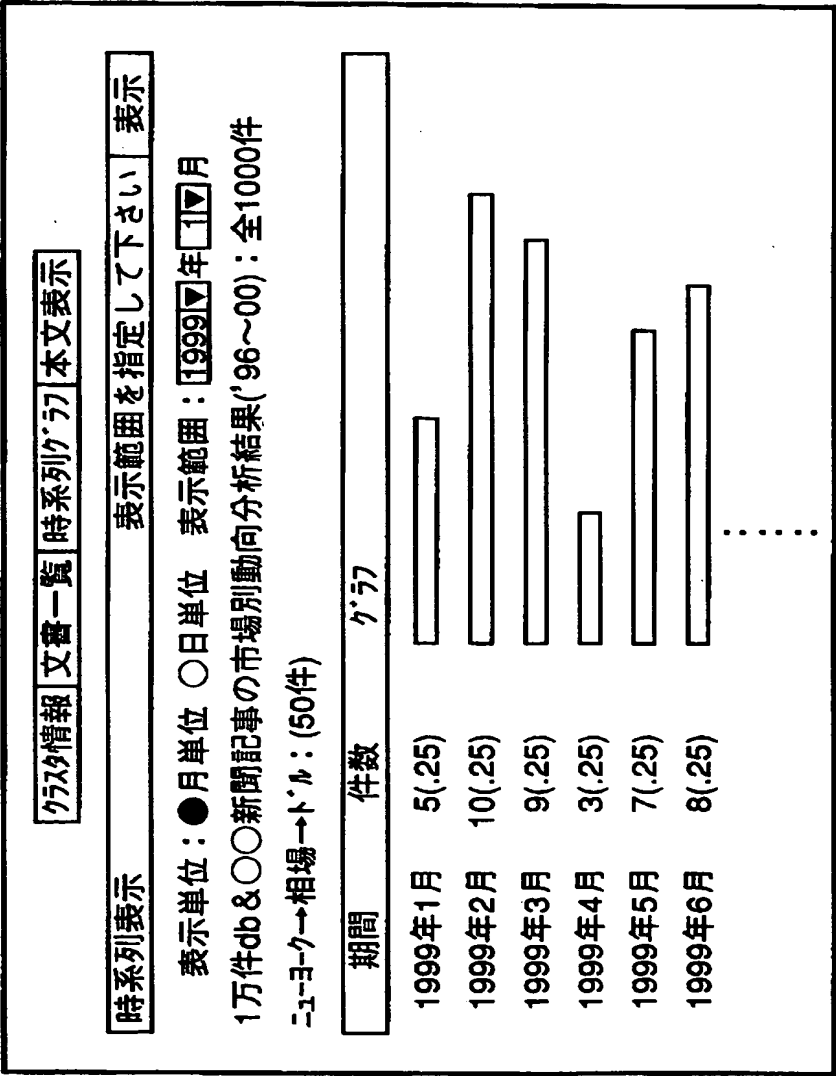
【図6】



【図7】



【図 8】



【図 9】

編集				
分析結果：5件		分析結果を選択して下さい		
分析結果	知識DB	件数	最上位 クラス数	更新日 作成日
<u>東京関連検索結果</u>	1万件db	1761件	10	2001/01/01 2001/01/01
<u>Faq2検索結果</u>	Faq2	222件	15	2001/02/02 2001/02/02
<u>エラーメッセージ</u>	ポータル 技術情報	333件	10	2001/03/03 2001/03/03
<u>ユーザーク関連</u>	1万件db, 1万件2	444件	10	2001/05/23 2001/04/04
<u>BIOS</u>	MAGNIA_FAQ	1000件	5	2001/05/10 2001/05/05
		...		

【図 10】

知識データベース：6件		分析対象の知識DBを選択して下さい		選択
<input type="checkbox"/>	対象	<input type="checkbox"/>	知識DB	コメント 件数
<input checked="" type="checkbox"/>				1万件db (No1 10225件)
<input checked="" type="checkbox"/>				〇〇新聞記事 (99年度 10000件)
<input type="checkbox"/>				Faq2 (Faq2 1234件)
<input type="checkbox"/>				MAGNIA_FAQ (MAGNIA_FAQ 5678件)
<input type="checkbox"/>				企業情報 (KYOCOSの企業情報を収集 555件)
<input type="checkbox"/>				特・ト技術情報 (特・ト技術情報 4810件)
			

【図11】

分析条件指定	分析条件を入力して下さい		実行
分析条件読込			
分析対象DB:	・1万件db(No1 10225件) ・朝日新聞記事(99年度 10000件)		
分析結果名称:	1万件db&〇〇新聞記事の市場別動向分析結果('96~00)		
分析対象期間:	1966年	1月	1日 ~ 2000年 12月 31日
絞込キーワード:	東京 ニューヨーク ロンドン		
絞込件数:	1000	件	
階層数:	〇1階層 ●多階層		
知識の重複:	●あり 〇なし		
最上位クラスの最大個数:	20	個	
重要語:	相場 市場 株 債券 年金 経済 銀行 金利 通貨		
不要語:	スポーツ 社会 事件		
同義語:	NY=NewYork=ニューヨーク; ●●●=○○○○;		

【図 1 2】

分析条件保存	保存する分析条件名称を入力して下さい	実行
分析条件名:	1万件db&〇〇新聞記事の市場別動向分析結果('96~00)	
分析対象期間:	1996年1月1日~2000年12月31日	
絞込キーワード:	東京 ニューヨーク ロンドン	
絞込件数:	1000件	
階層数:	多階層	
最上位クラスの最大個数:	20個	
文章の重複:	あり	
重要語:	相場 市場 株 債券 年金 経済 銀行 金利 通貨	
不要語:	スポーツ 社会 事件	
同義語:	NY=NewYork ニューヨーク ●●●=○○○	

【図 13】

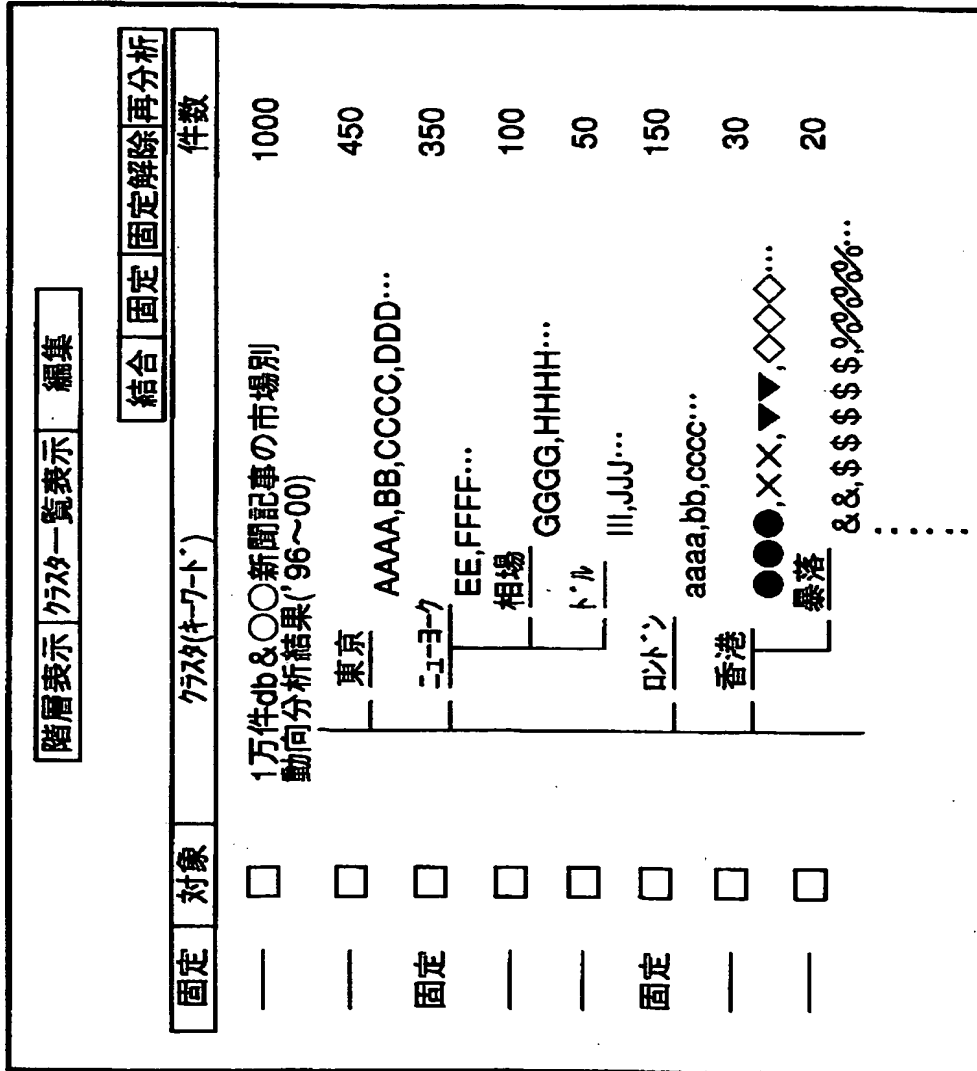
分析条件：全4件

使用する分析条件を選択して下さい

選択

対象	分析条件	作成日時
●	1万件db&〇〇新聞記事の市場別 動向分析結果	2001/01/01 00:23
○	▽▽▽DBの◆◆◆分析用設定	2001/02/02 09:20
○	CCCC	2001/03/03 18:00
○	DDDD	2001/04/01 22:00
	

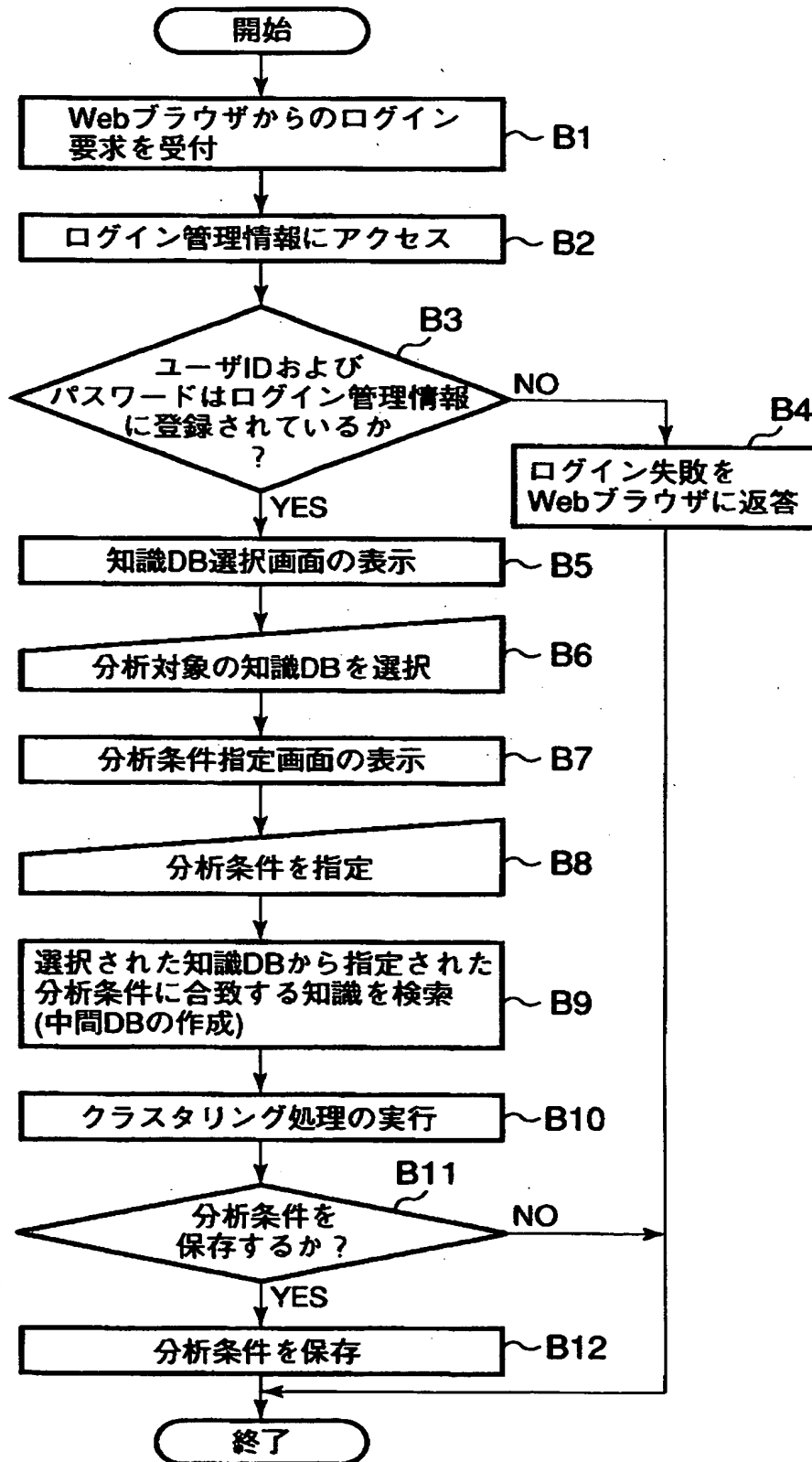
【図14】



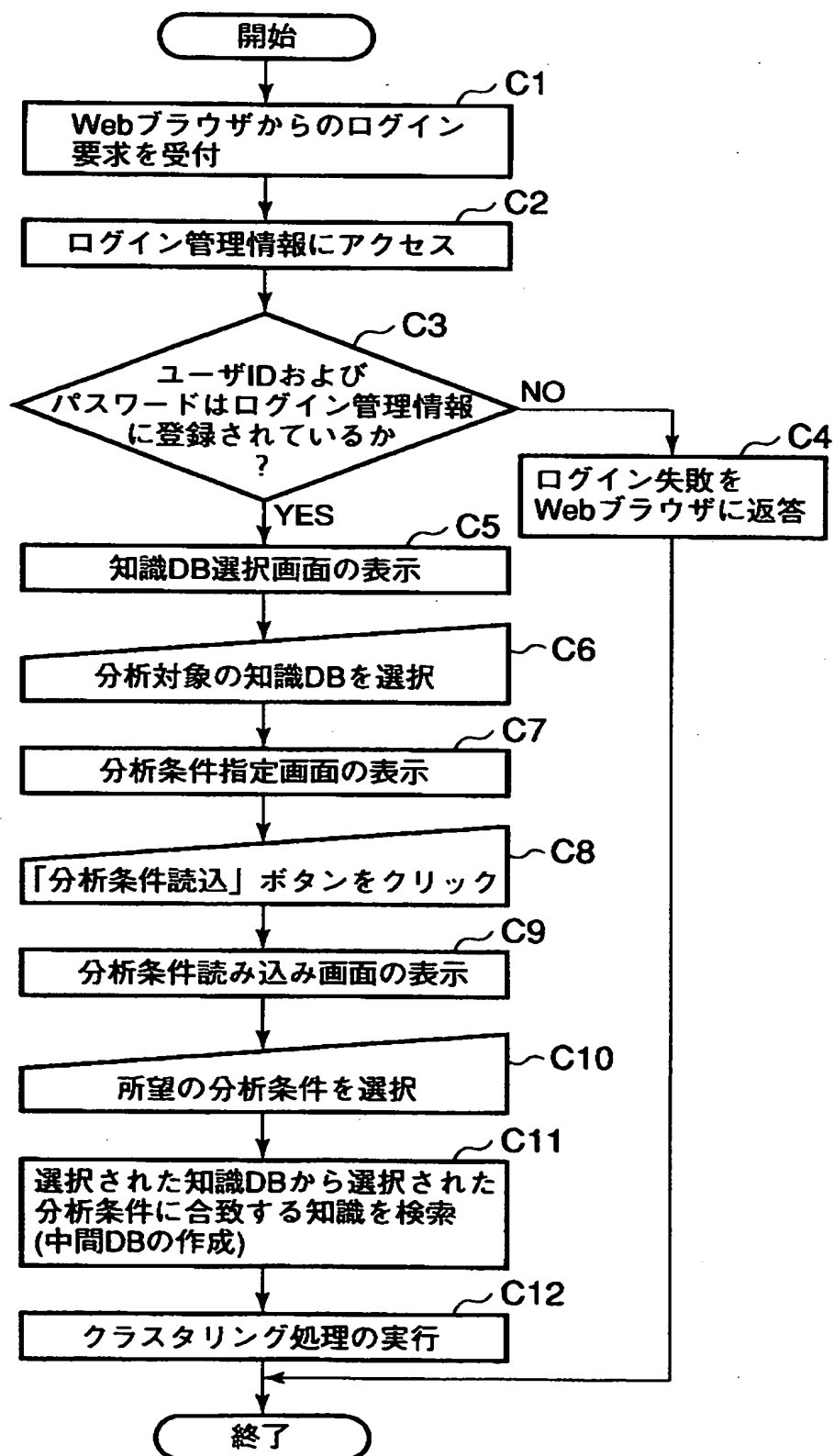
【図 15】

再分析条件指定	分析条件を入力して下さい	実行
分析対象DB：・1万件db(No1/10225件) ・〇〇新聞記事 99年度(10000件)		
分析結果名称：1万件db&〇〇新聞記事の市場別動向分析結果('96~00)		
分析対象期間：1996年1月1日~2000年12月31日		
絞込キーワード：東京 ニュヨーク ロンドン		
絞込件数：1000件		
階層数：多階層		
文章の重複：あり		
最上位クラスの最大個数： <input type="text" value="20"/> 個		
重要語：	<div>相場 市場 株 債券 年金 経済 銀行 金利 通貨 金融緩和 不良債券</div>	
不要語：	<div>スポーツ 社会 事件</div>	
同義語：	<div>NY=NewYork=ニユーヨーク; ●●●=○○○○;</div>	

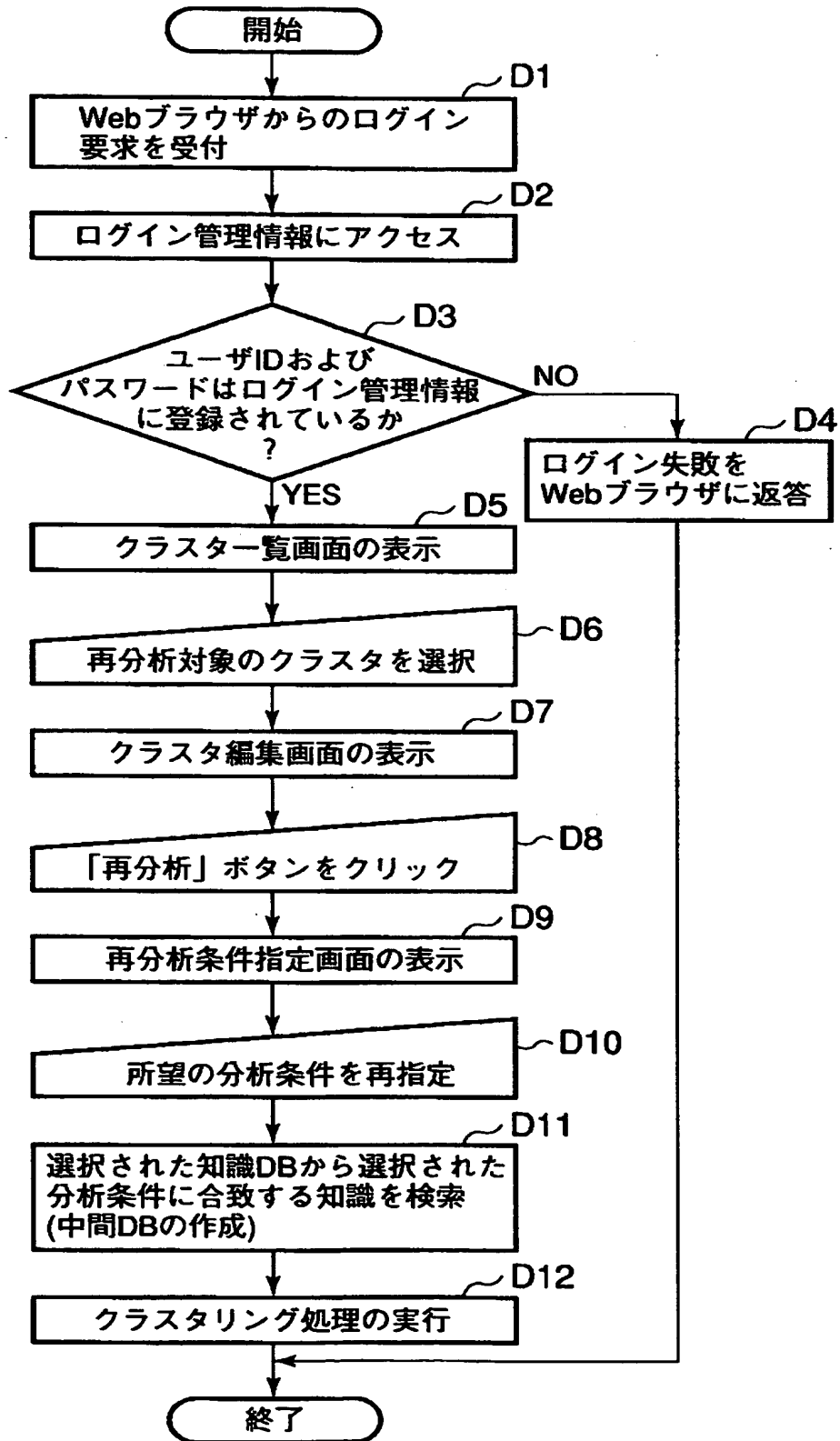
【図16】



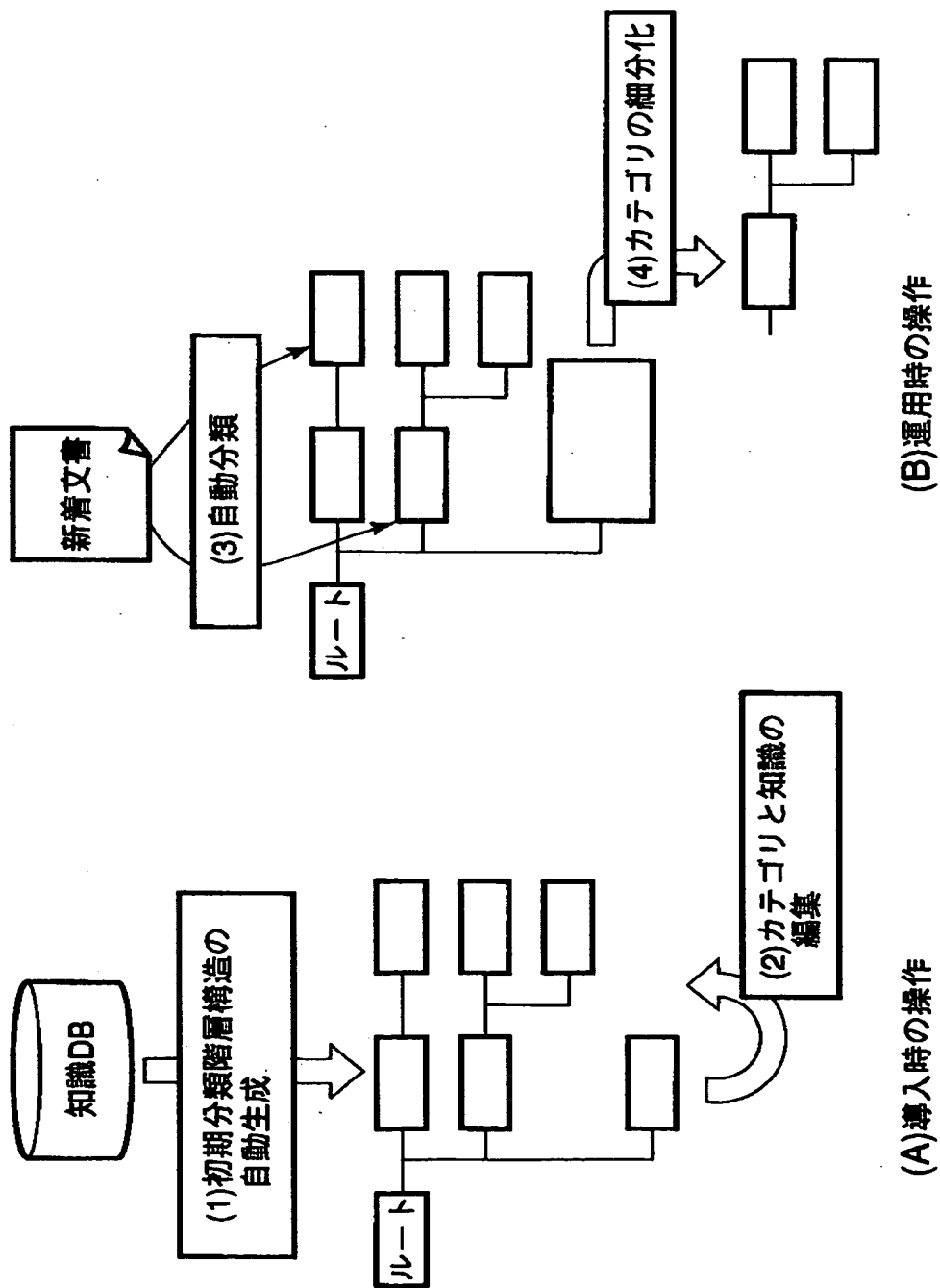
【図 17】



【図18】



【図 19】



【書類名】 要約書

【要約】

【課題】 クラスタの軸となる重要語およびこの重要語と組み合わせてクラスタリングを実行させるための不要語および同義語を設定可能とした知識分析システムを提供する。

【解決手段】 ナレッジサーバ122は、クライアント端末11から文書データベース1224に対して登録依頼された文書を別途インデックス化した知識として知識データベース1225に蓄積し、この知識データベース1225に蓄積した知識群から読み取れる傾向等の把握を支援するものである。そして、ナレッジサーバ122は、蓄積された知識群の内容的な偏りに左右されずに、ユーザの意図した通りの分類を実現するために、クラスタの軸となる重要語や不要語、同義語等をユーザに設定させる仕組みをもち、その設定に基づき、蓄積された知識群を知識クラスタに仕分けるいわゆるクラスタリング処理を実行して、その結果を分析結果データベース1227に格納する。

【選択図】 図1

出 願 人 履 歴 情 報

識別番号 [000003078]

1. 変更年月日	2001年 7月 2日
[変更理由]	住所変更
住 所	東京都港区芝浦一丁目1番1号
氏 名	株式会社東芝

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.